

Práctico 5 - MDP

Usted se encuentra en una aldea y quiere llegar a una isla. Sin embargo, en el medio puede encontrarse con un volcán que no le permite continuar el periplo. Para ello posee un mapa en donde está marcado el origen y todos los posibles estados finales marcados en verde o rojo. Cada celda es un estado. Desde cada estado se pueden tomar 4 acciones: N (norte), E (este), S (sur) y W (oeste).

		-50	20
Inicio		-50	
2			

En cada estado **s** si se toma la acción **a** se llegará al estado esperado con una probabilidad de $1 - slip_prob^1$ y al resto de estados con una probabilidad de $slip_prob$.

Se parte del estado **21**. El objetivo es llegar hasta el estado **14** que tiene una recompensa de 20, mientras que los estados **13** y **23** corresponden a estados finales con recompensas de -50. Otro estado final es el **31** pero tiene una recompensa de 2.

El mapa con estados se ve de la siguiente forma:

11	12	13	14
21	22	23	24
31	32	33	34

Se pide:

- 1) Calcular el valor de las estrategias N, S, E y W y compararlas.
- 2) Realizar un paso de Policy Improvement y analizar los cambios.
- 3) Realizar Policy Iteration y visualizar cuál es la policy de cada estado.

¹ La idea es ir modificando dicha variable y analizar los resultados.